

## Physical capability outcomes after total disc replacement with ProDisc-L

Humbert G. Sullivan, Robert L. Bobenmoyer, Kevin M. Boland, Molly M. Cerniglia, Vicki L. McHugh, Hayley L. Born, Michelle A. Mathiason and Nicholas R. Ladwig

*Int J Spine Surg* 2012, 6 () 43-48

doi: <https://doi.org/10.1016/j.ijsp.2011.11.001>

<https://www.ijssurgery.com/content/6/43>

This information is current as of September 15, 2024.

---

**Email Alerts** Receive free email-alerts when new articles cite this article. Sign up at:  
<http://ijssurgery.com/alerts>

## Physical capability outcomes after total disc replacement with ProDisc-L

Humbert G. Sullivan, MD <sup>a</sup>, Robert L. Bobenmoyer, OT <sup>b</sup>, Kevin M. Boland, MD <sup>c</sup>,  
Molly M. Cerniglia, PA-C <sup>a</sup>, Vicki L. McHugh, MS <sup>d,\*</sup>, Hayley L. Born, BS <sup>d</sup>,  
Michelle A. Mathiason, MS <sup>d</sup>, Nicholas R. Ladwig, BS <sup>d</sup>

<sup>a</sup> Department of Neurosurgery, Gundersen Lutheran Health System, La Crosse, WI

<sup>b</sup> Department of Industrial Rehabilitation, Gundersen Lutheran Health System, La Crosse, WI

<sup>c</sup> Department of Neuropsychology, Gundersen Lutheran Health System, La Crosse, WI

<sup>d</sup> Department of Medical Research, Gundersen Lutheran Medical Foundation, La Crosse, WI

### Abstract

**Background:** Lumbar disc arthroplasty (total disc replacement [TDR]) outcomes have been evaluated using subjective, patient-reported measures of pain, health, and functional impairment. As a condition of TDR coverage, our institution's health plan required that objective physical performance data be collected. Thus our study was designed to explore (1) the feasibility of using preoperative and 1-year postoperative performance on functional capacity tasks as an outcome metric for TDR with ProDisc-L (PD-L) (Synthes Spine, West Chester, Pennsylvania), (2) the magnitude and significance of changes in preoperative and postoperative performance, and (3) whether changes noted in performance are reflected in the subjective measures.

**Methods:** Seven adapted WorkWell tasks (physical capability assessment tool [PCAT]) (WorkWell Systems, Duluth, Minnesota) were performed preoperatively and 1 year postoperatively by 18 patients who received either single-level or 2-level PD-L implants. Demographic and medical data were reviewed.

**Results:** The PCAT was implemented easily, and the tasks took approximately 30 minutes to complete. Percent improvement and preoperative and postoperative physical capability outcomes for each PCAT task are as follows: squat, 79% (10.7 ± 7.1 repetitions vs 19.2 ± 2.0 repetitions,  $P < .001$ ); forward bend, 121% (110.2 ± 68.8 seconds vs 243.6 ± 77.2 seconds,  $P < .001$ ); kneel, 92% (283.2 ± 173.2 seconds vs 544.7 ± 109.3 seconds,  $P < .001$ ); floor-to-waist lift, 128% (16.1 ± 9.9 lb vs 36.7 ± 20.3 lb,  $P < .001$ ); horizontal carry, 119% (19.7 ± 8.6 lb vs 43.2 ± 18.3 lb,  $P < .001$ ); push, 32% (67.7 ± 19.2 lb vs 89.4 ± 24.4 lb,  $P < .001$ ); and pull, 40% (57.6 ± 17.1 lb vs 80.9 ± 26.4 lb,  $P < .001$ ). Visual analog scale scores for pain (5.1 ± 1.7 vs 1.4 ± 1.6,  $P < .001$ ), Oswestry Disability Index scores (49.0% ± 13.2% vs 15.2% ± 14.3%,  $P < .001$ ), and amount of narcotic use (26.1 ± 43.8 mg of morphine equivalent vs 1.9 ± 7.3 mg of morphine equivalent,  $P = .031$ ) also improved. In single-level cases, comparison of L4-5 versus L5-S1 showed significant differences only with the forward bend task ( $P = .002$ ).

**Conclusions/Clinical Relevance:** The physical capability outcome may be a feasible outcome metric. PD-L implantation may result in substantial improvements in physical performance. Similar benefits shown in a larger series over a longer timeframe could have important implications for the long-term health, productivity, and cost of health care for this patient population.

© 2012 ISASS - International Society for the Advancement of Spine Surgery. Published by Elsevier Inc. All rights reserved.

**Keywords:** ProDisc-L; Disc arthroplasty; Total disc replacement; Outcome assessment

Standard patient-reported outcome metrics used in spinal surgery (Oswestry Disability Index [ODI], Short Form 36, visual analog scale [VAS] score for pain, patient satisfac-

tion, and so on) provide important insight into a patient's view of personal health status and physical capabilities but do not allow one to objectively quantify or directly measure the changes in physical ability after surgery.<sup>1-4</sup> Self-reported outcome metrics may be influenced by factors such as the patient's poor recollection of prior events, a change in the patient's life situation, the patient's emotional state on the day of testing, or the patient's expectations of improved function after an intervention.<sup>1,5-7</sup> Concern from the health plan administration at our institution about the effect of these biases on self-reported outcomes led us to seek a more

Note: To access the supplementary table accompanying this report, visit the [Month] issue of *SAS Journal* at [journals.elsevierhealth.com/periodicals/ijsp/](http://journals.elsevierhealth.com/periodicals/ijsp/).

\* Corresponding author: Vicki L. McHugh, MS, Gundersen Lutheran Medical Foundation, 1900 South Ave, Mail Stop C03-006B, La Crosse, WI 54601; Tel: 608-775-3857; Fax: 608-775-1565.

E-mail address: [vlmchugh@gundluth.org](mailto:vlmchugh@gundluth.org)

2211-4599 © 2012 ISASS - International Society for the Advancement of Spine Surgery. Published by Elsevier Inc. All rights reserved.

10.1016/j.ijsp.2011.11.001

Downloaded from <https://www.ijssurgery.com/> by guest on September 15, 2024

objective outcome metric related to measurable physical performance. We report the results of our study using a new physical capability assessment tool (PCAT) in 18 consecutive subjects with ProDisc-L (PD-L) implantation (Synthes Spine, West Chester, Pennsylvania), comprising single- and 2-level implantation patients who had complete presurgical and postsurgical data.

The PCAT allows direct measurement of changes in physical performance after an intervention and was based on tasks validated for repeatability and relevance to real-world demands in the assessment of return to work and disability from a modified WorkWell workers' compensation functional capacity evaluation (WorkWell Systems, Duluth, Minnesota).<sup>8</sup> The magnitude of change in pounds lifted, force applied, or time a task can be sustained was designated as a physical capability outcome (PCO) for a specific task. Although there are limited published reports on the validity and reliability of the 3 categories of perceived exertions used in the WorkWell workers' compensation functional capacity evaluation, WorkWell was adapted from the 4-category Isernhagen work systems functional capacity evaluation. There have been multiple studies on the validity<sup>9,10</sup> and reliability<sup>11,12</sup> of the use of the Isernhagen functional capacity evaluation in both healthy subjects and those with chronic low-back pain. In developing the PCAT, we followed the 3-step procedure of Gouttebauge et al.<sup>13</sup> in the selection of functional tests from any full functional capacity evaluation method that assesses physical work ability in subjects with musculoskeletal complaints and related functional limitations.

In addition, a meta-analysis of 34 studies found that although the ODI and pain disability index had high levels of validity and reliability to assess functional capacity, no single functional test had a high level of both validity and reliability. However, a combination of both questionnaires and functional tests was recommended as the best instrument to assess functional capacity.<sup>14</sup> Our hypothesis was that adding a quantitative measurement of physical capacity improvement to the already validated ODI and VAS pain score would add an important new dimension of objectively documented physical performance to the standard subjective measures reflection of patient-perceived health status. This study is the first step in exploring the validity of this hypothesis.

## Methods

In this institutional review board–approved retrospective study, all patients who had undergone total disc replacement (TDR) with PD-L between May 29, 2007, and June 17, 2009, and had completed preoperative and 1-year postoperative assessments met the criteria for study inclusion. This population included patients who underwent single-level, multilevel, or hybrid PD-L procedures. Patients' electronic and paper medical records were retrospectively reviewed, and the following data were collected: age; sex; body mass

index; smoking status; vertebral-level location of surgery; workers' compensation and litigation status; pain duration before surgery; preoperative and postoperative occupation; preoperative medical and neuropsychological evaluation results; results of discography; and preoperative and 1-year postoperative assessment of the ODI, VAS pain score, PCO, and narcotic use (expressed as milligrams of morphine equivalent for 24 hours, based on the average of 1 week's data, using the Web site <http://www.agencymeddirectors.wa.gov/files/dosingcalc.xls>).

To qualify for TDR with PD-L implantation, patients were required to meet all the Food and Drug Administration inclusion criteria for surgical treatment for PD-L.<sup>15</sup> Moreover, patients must have had low-back pain with or without leg pain for a minimum of 1 year, a bone mineral density T-score by dual-energy X-ray absorptiometry scan greater than or equal to  $-1.0$ , and at least a single level of degenerative disc disease at L3-4, L4-5, or L5-S1 on magnetic resonance imaging. Neuropsychological evaluation data following the presurgical psychological screening algorithm of Block<sup>16</sup> were used to exclude any patient receiving less than a good or fair surgical candidate rating.

Twenty-seven subjects met TDR inclusion criteria and underwent PD-L implantation during the defined study period. Of these subjects, 4 were excluded for undergoing a hybrid procedure that included 1 PD-L implant and 1 fusion, 1 was excluded for undergoing a 3-level PD-L implantation, and 3 were excluded because of incomplete data. In addition, 1 subject had bony injury 10 days postoperatively due to an all-terrain vehicle accident and was excluded. This left a cohort of 18 subjects who completed preoperative and 1-year follow-up visits and were included in the analysis.

### PCO assessment

The assessment tasks comprising the PCAT are standard, well-researched measures of known variance used in fit-for-work evaluations.<sup>10,11,17–20</sup> The elements included in the PCAT evolved through a series of trials over a 4-year period (2005–2008), initially using elements of the WorkWell workers' compensation functional capacity evaluation<sup>8</sup> to objectively measure change in physical capability of patients with chronic pain in response to various therapies. Specifically, we devised the assessment tool for TDR by adapting lumbar components of the WorkWell evaluation known to be altered in patients with low-back pain. Of the 26 components of the WorkWell evaluation, 7 have been adopted as the tasks composing the PCAT for the PCO evaluation.

The tasks composing the PCAT are as follows: (1) walk, measured in yards walked in 6 minutes (warm-up exercise only), with a standard mean of 535 yd; (2) squats, comprising up to 20 repetitions; (3) forward bend, for which the patient maintained a light board game task in the standing bent-forward position for up to 300 seconds; (4) kneel, for which the patient maintained a light board game task in the kneeling position for up to 600 seconds; (5) lift, defined as

a floor-to-waist lift, for 5 repetitions, measured in pounds; (6) carry, defined as a horizontal lift, for 50 feet, with 1 repetition, measured in pounds; and (7) push, defined as a static push force, and pull, defined as a static pull force, both measured in pounds.

An endpoint for each of the tasks was defined in 1 of the following ways: baseline patient-reported symptoms increased with or without kinesio-physical changes in performance, the kinesio-physical maximum was reached (first noted activation of accessory muscles with inability to maintain proper form), or the predetermined task maximum was reached. To limit the time required to perform the evaluation, task maxima for timed interventions were adopted from the WorkWell workers' compensation functional capacity evaluation.<sup>8</sup> A change in PCO is considered significant when the measured quantities before and after intervention are statistically different from one another for a cohort of subjects ( $P \leq .05$ ). A negative PCO is a deterioration in performance. A positive PCO is an improvement in performance. Magnitude and statistical significance of PCOs from the PCAT administered before and 1 year after surgery are reported. Qualitative comparison of the PCOs with the standard, patient-reported outcome measures is also provided to determine whether patient perception of outcomes supports the more objective results of the newly developed PCAT.

The PCAT was administered by a single designated occupational therapist. The push and pull tasks were measured with Chatillon force gauges (Ametek, Largo, Florida). A video and complete description of each of these tasks are available at <http://www.gundluth.org/PCO>.

### *Surgical technique*

The surgical technique used in the single-level PD-L implantation followed that described in the standard US technique for the PD-L investigational device exemption (IDE) study.<sup>21</sup> Two-level and hybrid procedures followed the same protocol at multiple levels. All patients were operated on while under general anesthesia.

### *Radiographic evaluation*

Synapse (version 3.1.1, Fujifilm Medical Systems, Stamford, Connecticut), a picture archiving and communication system and a Digital Imaging and Communications in Medicine-compliant viewing system, was used for angle measurement of flexion/extension between the 2 keels of the PD-L implant.

### *Statistical analysis*

Presurgical and postsurgical statistical comparisons, including PCO, ODI, VAS pain score, and use of morphine equivalents, were calculated with paired *t* tests. Results comparing single-level L4-5 cases with single-level L5-S1 cases were calculated by use of 2-sample *t* tests.  $P < .05$  was considered significant. The statistical analysis was con-

ducted with Microsoft Excel 2010 (version 14; Microsoft, Redmond, Washington).

## **Results**

Of the 27 subjects who met TDR inclusion criteria and underwent PD-L implantation, 18 met inclusion criteria and were included in the analysis. The first 17 PD-L implantation procedures were completed by a single surgeon, and all exposures were performed by 1 of 2 experienced access surgeons. The implantation for subject 18 was assisted by a visiting surgeon using a modified technique that incorporated the use of pilot holes parallel to the apex of the device keels.<sup>22</sup>

The 18 subjects (6 men and 12 women) had a mean age of  $40.0 \pm 10.2$  years (range, 25–62 years) at the time of surgery. The mean body mass index was  $25.7 \pm 3.2$  kg/m<sup>2</sup> (range, 20.7–30.9 kg/m<sup>2</sup>). At the time of surgery, 3 subjects (17%) (subjects 2, 5, and 9) were involved in litigation regarding their back injury and 5 subjects (28%) (subjects 2, 5, 7, 8, and 18) were receiving workers' compensation. Of the subjects, 6 (33%) (subjects 2, 5, 6, 8, 15, and 16) were smokers, but all claimed to have quit in preparation for surgery. Overall, age, overweight status, and smoking status were not associated with positive outcomes.

On discography, all operated discs were positive for reproduction of concordant pain, with the level above as a negative control. Neuropsychological testing following the presurgical psychological screening algorithm<sup>16</sup> showed that 94% of subjects were considered a good risk for surgery, with only 1 subject being regarded as a fair risk for surgery (subject 2). Of the 18 subjects, 15 (83%) underwent single-level TDR with PD-L implantation: 6 subjects underwent TDR at L4-5 (subjects 2, 3, 4, 6, 13, and 17) and 9 subjects at L5-S1 (subjects 1, 5, 9, 11, 12, 14, 15, 16, and 18). Three subjects received 2-level TDR with PD-L implantation at the L4-5 and L5-S1 levels (subjects 7, 8, and 10).

Administering the PCAT took approximately 30 minutes before and after surgery. Individual subject results are presented in [Supplementary Table 1](#) (online only, available at [journals.elsevierhealth.com/periodicals/ijsp/](http://journals.elsevierhealth.com/periodicals/ijsp/)). Overall mean results, summarized in [Table 1](#), show significant improvements in each PCO assessment task. [Figure 1A](#) illustrates the dramatic mean improvement in the forward bend and kneeling tasks, whereas [Fig. 1B](#) depicts mean improvements in the floor-to-waist lift, forward 50-yd carry, static push, and static pull tasks. There were also significant improvements in the ODI, VAS pain score, and narcotic use. Particularly noteworthy were the individual maximum improvements, not all in a single individual, of 1900% in squats (1 repetition before surgery vs 20 after surgery), 1011% in forward bend (27 seconds before surgery vs 300 seconds after surgery), and 700% in floor-to-waist lift (5 lb before surgery vs 25 lb after surgery). The maximum improvements in time and repetitions noted are an underesti-



Table 1

Summary of physical capability testing, ODI, VAS pain score, and pain medication use preoperatively and 12 months postoperatively: Mean data and percent change

PCO	No. of patients	Preoperative		Postoperative		Improvement		P value (paired <i>t</i> test)	% Improvement
		Mean	SD	Mean	SD	Mean	SD		
Squats (maximum, 20 repetitions) (repetitions)	18	10.7	7.1	19.2	2.0	8.5	7.3	< .001	79*
Forward bend (maximum, 300 s) (s)	18	110.2	68.6	243.6	77.2	133.4	86.5	< .001	121*
Kneel (maximum, 600 s) (s)	18	283.2	173.2	544.7	109.3	261.5	173.8	< .001	92*
Floor-to-waist lift (lb)	18	16.1	9.9	36.7	20.3	20.6	15.0	< .001	128
Forward 50-yd carry (lb)	17†	19.7	8.6	43.2	18.3	23.5	16.3	< .001	119
Push (lb)	17†	67.7	19.2	89.4	24.4	21.7	17.8	< .001	32
Pull (lb)	17†	57.6	17.1	80.9	26.4	23.4	24.4	< .001	40
ODI (%)	18	49.0	13.2	15.2	14.3	33.8	17.6	< .001	
VAS pain score	18	5.1	1.7	1.4	1.6	3.7	1.8	< .001	
Morphine equivalency	16†	26.1	43.8	1.9	7.3	24.3	39.5	.031	

NOTE. Because of rounding when the final calculation was performed, calculations based on preoperative and postoperative mean values may not equal the mean change or percent change values presented.

\* Percentages are not indicative of total improvement because patients reached the maximum value.

† Patients were unable to complete the task because of an unrelated injury or data were missed during collection.

mation, because several subjects attained the predetermined task maximum with ease. When we compared single-level cases by level (L4-5 vs L5-S1), the only statistically significant comparison was for the forward bend task ( $P = .002$ ) (Table 2). A similar analysis comparing single-level versus 2-level cases yielded no significant differences.

Correspondingly, the ODI scores changed from a range of 22% to 84% disability preoperatively to a range of 0% to 50% disability postoperatively. The mean 1-year lumbar intersegmental range of motion as determined by flexion-extension radiographs was 12° (range, 4°–25°), which verified that all subjects had functioning devices. Furthermore, at 1 year postoperatively, 16 of 18 subjects (89%) had ceased taking all narcotic medications. No patient had clinically significant complications related to surgery.

## Discussion

The objectives of this pilot study were 3-fold: to evaluate the feasibility of adapting tasks from WorkWell to objectively and quantitatively evaluate changes in physical capability after PD-L implantation; to evaluate whether the magnitude of change in physical capability is statistically or clinically significant and whether this testing should be further evaluated in a larger cohort of patients; and to assess whether objective, quantitative physical capacity measurement qualitatively parallels results of the standard self-reported measures.

Feasibility was shown by the ease of implementation of the PCOs in the clinical setting, requiring only approximately 30 minutes per evaluation by a skilled functional capacity evaluator. The statistically and clinically significant improvement in physical capacity noted in all but 1 of the PCO tasks at 1 year after PD-L surgery provided an objective, quantitative complement to the self-reported outcomes and reduction in narcotic use. As seen in the PCAT,

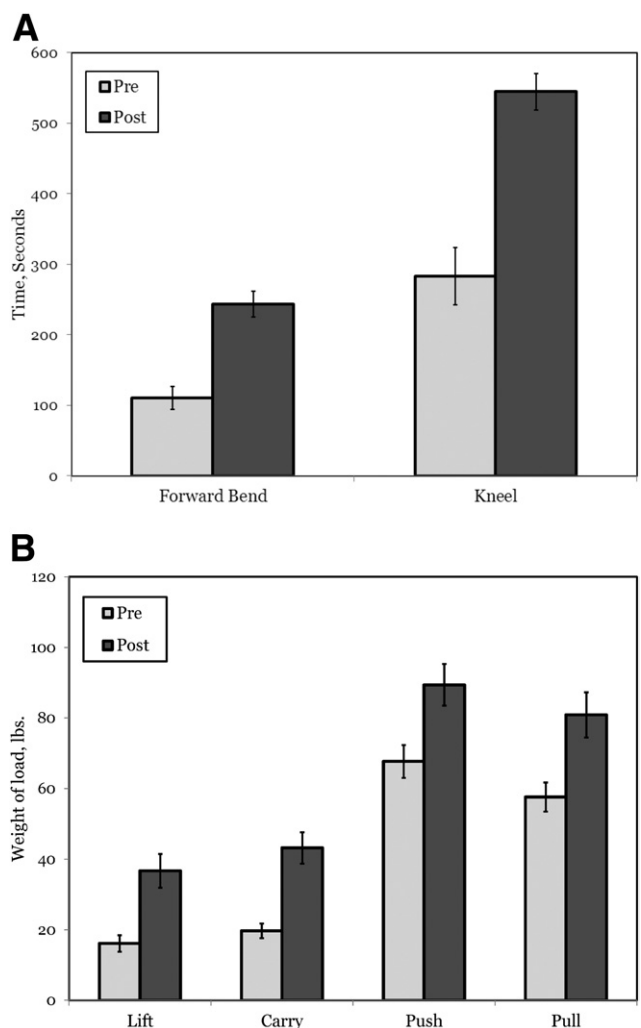


Fig. 1. Overall mean presurgical (Pre) and postsurgical (Post) performance for forward bend and kneeling tasks (A) and floor-to-waist lift, forward 50-yd carry, static push, and static pull tasks (B). Error bars indicate the standard error of the mean.

Table 2

Summary of physical capability testing, ODI, VAS pain score, and pain medication use for single-level L4-5 versus L5-S1: Mean presurgical versus postsurgical data and percent change

PCO	No. of patients	Preoperative		Postoperative		Improvement		% Improvement
		Mean	SD	Mean	SD	Mean	SD	
<b>L4-5</b>								
Squats (maximum, 20 repetitions) (repetitions)	6	11.2	7.2	18.8	2.9	7.7	6.5	68*
Forward bend (maximum, 300 s) (s)	6	127.3	94.2	178.3	101.7	51.0	66.2	40*
Kneel (maximum, 600 s) (s)	6	163.0	30.2	466.8	169.9	303.8	170.6	186*
Floor-to-waist lift (lb)	6	17.5	11.7	35.8	28.4	18.3	19.9	105
Forward 50-yd carry (lb)	6	23.3	12.9	37.5	21.9	14.2	15.3	61
Push (lb)	6	65.3	20.5	87.7	25.2	22.3	11.8	34
Pull (lb)	6	56.8	20.8	88.3	37.2	31.5	35.2	55
ODI (%)	6	59.0	15.6	18.0	16.6	41.0	23.6	
VAS pain score	6	4.5	1.2	1.3	1.6	3.2	2.0	
Morphine equivalency	5†	4.0	8.9	0.0	0.0	4.0	8.9	
<b>L5-S1</b>								
Squats (maximum, 20 repetitions) (repetitions)	9	12.3	7.8	19.2	2.0	6.9	8.2	56*
Forward bend (maximum, 300 s) (s)	9	107.3	62.3	284.1	33.8	176.8	60.3	165*
Kneel (maximum, 600 s) (s)	9	335.2	199.1	578.2	43.4	243.0	192.4	72*
Floor-to-waist lift (lb)	9	18.9	8.9	43.3	15.2	24.4	14.2	129
Forward 50-yd carry (lb)	9	18.9	6.0	50.6	16.3	31.7	16.2	168
Push (lb)	9	70.3	22.4	94.8	27.9	24.4	22.9	35
Pull (lb)	9	59.3	18.4	80.9	23.3	21.6	19.1	36
ODI (%)	9	44.7	6.8	15.2	15.9	29.4	15.0	
VAS pain score	9	5.6	1.8	1.6	2.0	4.0	1.9	
Morphine equivalency	9†	44.2	54.3	3.3	10.0	40.9	48.7	

NOTE. Because of rounding when the final calculation was performed, calculations based on preoperative and postoperative mean values may not equal the mean change or percent change values presented.

\* Percentages are not indicative of total improvement because patients reached the maximum value.

† Patients were unable to complete the task because of an unrelated injury or data were missed during collection.

all changes were statistically and clinically relevant and carried over to real-world tasks. For instance, an increase of 20 lb in the floor-to-waist lift, forward 50-yd carry, push, and pull variables was shown in many patients. In addition, the improvement seen in ODI scores in our study compares favorably with the IDE study.<sup>20</sup>

Although previous studies have shown the superiority and/or equal efficacy of PD-L implantation compared with fusion in a number of categories,<sup>20,21,23</sup> it remains a controversial option for patients with severe, intractable low-back pain due to degenerative disc disease.<sup>24</sup> One issue adding to the paradox has been the use of subjective outcome measures in IDE studies, consisting of patient reporting of health status, functional impairment, and pain.<sup>15</sup> We believe we have developed a promising objective outcome measure for the value of TDR that may also serve as a model for comparable objective outcome assessments evaluating other types of procedures.

We postulate that this type of PCO assessment may have potential use in future IDE studies. In this setting, the more objective nature of the PCO assessment, compared with the commonly used self-reported measures, may reduce the effect of expectation biases related to randomization (the expectation of a good result by patients receiving the trial device and the concern about a poor outcome by patients in the control arm). No study has attempted to directly measure the expectation biases. One indirect attempt did not account for

the positive expectation bias in the continuing-access arm of the PD-L IDE study (the patients were receiving something not generally available).<sup>25</sup> If larger cohort studies continue to suggest the usefulness of PCO assessment, then perhaps PCO assessment and expectation bias evaluation could become part of future IDE testing.

The major limitation of this study is its small sample size. In addition, the use of 3 patient-reported outcome measures commonly used in outcome studies limits comparisons to only these measurements (ODI, VAS pain score, and narcotic use). Furthermore, future studies should note when determination of PCO endpoints has occurred only in relation to patient-reported exacerbation of physical symptoms without kinesiophysical changes. This is important because it detracts from the objectivity of the evaluations. Our impression is that this endpoint was infrequent at 1 year and common preoperatively, but it was not specifically recorded in our pilot patients.

Each patient's performance on physical capability tasks can be related to whether that performance is sufficient for his or her presurgical occupation, although this was not measured in our study. The number of tasks performed at a level appropriate for the patient's occupation could then be compared between the preoperative and postoperative measurements. Finally, it should be emphasized that our results represent pilot data, and a larger study would allow for a

more rigorous statistical analysis to validate the results in a larger cohort of arthroplasty and fusion patients.

## Conclusion

The combination of objective measurement of changes in physical capacity and the standard subjective outcome measures may provide a more complete picture of a patient's response to surgery than either measure alone. These data suggest that use of PCO assessment in combination with standard subjective outcome measures may be a promising assessment tool and merits further study in a larger sample.

On the basis of the conclusion of this study, a more extensive battery of patient-reported measures combined with PCO assessment in a larger cohort seems to be indicated and has been incorporated in ongoing studies. Additional studies will allow for a more rigorous statistical analysis to validate the results in a larger cohort of both arthroplasty and fusion patients.

## Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.ijsp.2011.11.001](https://doi.org/10.1016/j.ijsp.2011.11.001).

## References

- Smith DM, Brown SL, Ubel PA. Mispredictions and misrecollections: challenges for subjective outcome measurement. *Disabil Rehabil* 2008;30:418–24.
- Khorsan R, Coulter ID, Hawk C, Choate CG. Measures in chiropractic research: choosing patient-based outcome assessments. *J Manipulative Physiol Ther* 2008;31:355–75.
- Flaherty SA. Pain measurement tools for clinical practice and research. *AANA J* 1996;64:133–40.
- Greenhalgh J, Meadows K. The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: a literature review. *J Eval Clin Pract* 1999;5:401–16.
- McCormack HM, Horne DJ, Sheather S. Clinical applications of visual analogue scales: a critical review. *Psychol Med* 1988;18:1007–19.
- Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analogue scale. *Pain* 1983;16:87–101.
- Koller M, Hoffmann S, Rothmund M, Lorenz W, Plaul U. Checking for interviewer bias in outcome assessment: a method for strengthening the design of prospective, randomised trials in surgery. *Langenbecks Arch Surg* 2004;389:92–6.
- WorkWell Systems I: *Functional Capacity Evaluation*. Duluth, MN: WorkWell Systems; 2006.
- Reneman MF, Fokkens AS, Dijkstra PU, Geertzen JH, Groothoff JW. Testing lifting capacity: validity of determining effort level by means of observation. *Spine (Phila Pa 1976)* 2005;30:E40–6.
- Reneman MF, Joling CI, Soer EL, Göeken LN. Functional capacity evaluation: ecological validity of three static endurance tests. *Work* 2001;16:227–34.
- Reneman MF, Brouwer S, Meinema A, Dijkstra PU, Geertzen JH, Groothoff JW. Test-retest reliability of the Isernhagen work systems functional capacity evaluation in healthy adults. *J Occup Rehabil* 2004;14:295–305.
- Reneman MF, Jaegers SM, Westmaas M, Göeken LN. The reliability of determining effort level of lifting and carrying in a functional capacity evaluation. *Work* 2002;18:23–7.
- Gouttebauge V, Wind H, Kuijper PP, Sluiter JK, Frings-Dresen MH. How to assess physical work-ability with Functional Capacity Evaluation methods in a more specific and efficient way? *Work* 2010;37:111–5.
- Wind H, Gouttebauge V, Kuijper PP, Frings-Dresen MH. Assessment of functional capacity of the musculoskeletal system in the context of work, daily living, and sport: a systematic review. *J Occup Rehabil* 2005;15:253–72.
- Andersen DG, Vaccaro AR, Tannoury C. Patient selection for fusion vs. motion-sparing surgery for symptomatic degenerative disk disease. In: Vaccaro AR, Papadopoulos S, Traynelis V, Haid RJ, Sasso R, eds. *Spinal Arthroplasty: The Preservation of Motion*. Edinburgh: Elsevier Saunders; 2007:11–20.
- Block AR. *Presurgical Psychological Screening in Chronic Pain Syndromes: A Guide for the Behavioral Health Practitioner*. Mahwah, NJ: Lawrence Erlbaum Associates; 1996.
- Reneman MF. *Functional Capacity Evaluation in Patients With Chronic Low Back Pain: Reliability and Validity* [dissertation]. Groningen: University Library Groningen; 2004. Available from: URL: <http://dissertations.ub.rug.nl/faculties/medicine/2004/m.f.reneman/>. Accessed November 10, 2011.
- Gross DP, Battié MC. Reliability of safe maximum lifting determinations of a functional capacity evaluation. *Phys Ther* 2002;82:364–71.
- Isernhagen SJ, Hart DL, Matheson LM. Reliability of independent observer judgments of level of lift effort in a kinesio-physical functional capacity evaluation. *Work* 1999;12:145–50.
- Zigler J, Delamarter R, Spivak JM, et al. Results of the prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement versus circumferential fusion for the treatment of 1-level degenerative disc disease. *Spine* 2007;32:1155–62, discussion 1163.
- Delamarter R, Zigler JE, Balderston RA, Cammisa FP, Goldstein JA, Spivak JM. Prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement compared with circumferential arthrodesis for the treatment of two-level lumbar degenerative disc disease: results at twenty-four months. *J Bone Joint Surg Am* 2011;93:705–15.
- Sullivan HG, Bertagnoli R, Ladwig NR, et al. Prevention of vertebral body vertebral body-splitting fractures following multi-level ProDisc-L implantation. *SAS J* In press 2012.
- Lin EL, Wang JC. Total disk arthroplasty. *J Am Acad Orthop Surg* 2006;14:705–14.
- Whang PG, Simpson AK, Rehtine G, Grauer JN. Current trends in spinal arthroplasty: an assessment of surgeon practices and attitudes regarding cervical and lumbar disk replacement. *J Spinal Disord Tech* 2009;22:26–33.
- Zigler J, Ohnmeiss D. Is there a euphoric bias affecting the outcomes of randomized vs. nonrandomized patients enrolling in new technology trials? [abstract]. *Spine J* 2008;8(Suppl 1):58S–9S. Available from: URL: <http://www.sciencedirect.com/science/article/pii/S1529943008003987>. Accessed November 10, 2011.